ELSEVIER

Stochastic Processes and their Applications 53 (1994) 193–220

stochastic processes and their applications

# A long range sexual reproduction process

Claudia Neuhauser*, [1]

*Department of Mathematics, University of Wisconsin, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706, USA*

Received 21 July 1992; revised 29 November 1993

**Abstract**

We describe a long range growth model with sexual reproduction on $\varepsilon\mathbb{Z}$ in which particles die at rate 1 and pairs of adjacent particles produce offspring at rate $\lambda$. The offspring is sent to a site chosen at random from the neighborhood of the parent particles. If the site is already occupied, the birth is suppressed, that is, we allow at most one particle per site. The size of the neighborhood increases as $\varepsilon$ tends to 0. We investigate the behavior for small $\varepsilon$. In the limit as $\varepsilon \to 0$, the particle density evolves according to an integro-differential equation which has traveling wave solutions whose wave speed is a nondecreasing function of $\lambda$. We compare the system for small $\varepsilon$ with the limiting system and discuss phase transition. We show that the behavior of the particle system for sufficiently small $\varepsilon$ is similar to the behavior of the limiting system. That is, if $\lambda$ is sufficiently small so that the wave speed of the traveling wave of the limiting equation is negative, then, for small enough $\varepsilon$, the pointmass at the all-empty configuration is the only stable equilibrium. If $\lambda$ is large enough, so that the wave speed of the traveling wave of the limiting equation is positive, then, for small enough $\varepsilon$, the system has a positive probability of survival, that is, in addition to the pointmass at the all-empty configuration, there is a nontrivial equilibrium. Not predicted by the limiting system, there is a range of values of $\lambda$ for which the all-empty configuration is the only stable equilibrium but for which the particle system exhibits metastable behavior.

*Keywords:* Interacting particle systems; Growth model; Long range process; Metastability; Phase transition; Sexual reproduction process

## 1. Introduction

We study a simple growth model on a discrete one-dimensional lattice known as the *contact process with sexual reproduction.* In this model, particles die and adjacent pairs of particles give birth to new particles which are then sent to a site in the neighborhood set of the parent particles. If a particle is sent to an already occupied site, the birth is suppressed. That is, we allow at most one particle per site. We

* Corresponding author.
[1] Partially supported by NSF grant DMS-9206139 and MSI.


SSDI 0304-4149(93)E0099-Z

investigate the case in which the neighborhood set is very large. We model this system as a continuous time Markov process in which the state at time $t$ is $\xi_t^\varepsilon : \varepsilon\mathbb{Z} \to \{0,1\}$. If $\xi_t^\varepsilon(x) = 0$ for some $x \in \varepsilon\mathbb{Z}$, then we say $x$ is vacant; otherwise we say that $x$ is occupied by a particle. We also use set notation and say $x \in \xi_t^\varepsilon$ if and only if $\xi_t^\varepsilon(x) = 1$. We rescale the integer lattice $\mathbb{Z}$ by a scale parameter $\varepsilon > 0$. Small $\varepsilon$ corresponds to a large neighborhood set. The dynamics are given by:

(i) 1's die (i.e., become 0's) at rate 1.
(ii) Pairs of particles in adjacent sites produce offspring at rate $\lambda$.
(iii) If the parents of a particle are at $x$ and $x + \varepsilon$, the offspring is sent to a site $y$ chosen at random according to a probability kernel $k_\varepsilon(x - y)$. (We will specify $k_\varepsilon(x - y)$ below.)
(iv) If $y$ is already occupied the birth is suppressed.

Since the process does not have spontaneous births, $\delta_\emptyset$, the pointmass at the all-empty configuration, is a stationary distribution. $\delta_\emptyset$ is called the trivial equilibrium. Interest focuses on the question of whether there are nontrivial equilibria in which $\lim_{t\to\infty} P(\xi_t^\varepsilon(0) = 1) > 0$.

It is easy to see that the system is attractive. That is, if $\xi_0^\varepsilon \le \eta_0^\varepsilon$ (i.e., $\xi_0^\varepsilon(x) \le \eta_0^\varepsilon(x)$ for all $x \in \varepsilon\mathbb{Z}$), then we can construct copies of the process with these initial configurations so that $P(\xi_t^\varepsilon \le \eta_t^\varepsilon) = 1$ for all $t$. Let $\xi_t^{\varepsilon,1}$ be the configuration at time $t$ when initially $\xi_0^{\varepsilon,1}(x) \equiv 1$ for all $x \in \varepsilon\mathbb{Z}$. Then as a consequence of attractiveness, $\xi_t^{\varepsilon,1}$ will converge weakly to the largest invariant measure $\xi_\infty^{\varepsilon,1}$. If $\xi_\infty^{\varepsilon,1} = \delta_\emptyset$, then $\delta_\emptyset$ is the only invariant distribution and we say that the system dies out. If they are different, we say the system has a positive probability of survival. Since the system is attractive, it follows from Theorem 2.14 in Chapter III of Liggett (1985) that there exists a critical value $\lambda_c$ such that $\delta_\emptyset = \xi_\infty^{\varepsilon,1}$ if $\lambda < \lambda_c$ and $\delta_\emptyset \ne \xi_\infty^{\varepsilon,1}$ if $\lambda > \lambda_c$. That is,

$$\lambda_c = \inf\{\lambda : P(\xi_\infty^{\varepsilon,1}(0) = 1) > 0\}.$$

It is easy to see that $\lambda_c \ge 1$ by comparison with the asexual contact process. (In the asexual contact process only one parent is needed to produce an offspring. The death mechanism is the same as in the sexual contact process. For more on the asexual contact process see Liggett (1985) or Durrett (1988).) Durrett and Gray (1986) showed that $\lambda_c < \infty$.

Before we state our results, we will give a brief history of the process. A two-dimensional version with an asymmetric neighborhood (that is, the parents of an offspring at $x$ are located at $x + (1,0)$ and $x + (0,1)$) was studied by Durrett and Gray. A summary of their results can be found in Durrett and Gray (1986). Chen (1992) investigated the stability of the absorbing state $\delta_\emptyset$ for the two-dimensional system with nearest neighbor interaction. (That is, the two parents of an offspring at $x$ are located at sites chosen at random from the four nearest neighbors of $x$). One of his results was that for any given $\lambda \in (1, \infty)$ if the initial state is a product measure with sufficiently low density, then the system will die out. We will see that this no longer holds true for the one-dimensional system when the range of interaction is sufficiently large.

In this paper, we investigate the one-dimensional system for small $\varepsilon$. In order to get any results it is necessary to make some assumptions on the kernel $k_\varepsilon(x)$ mentioned in (iii). $k_\varepsilon(x)$ is derived from a kernel $k(x)$ for which we assume throughout the paper that the following conditions hold:

(K1) $k$ is a piecewise continuous, nonnegative, and even function on $\mathbb{R}$ with $\int_{\mathbb{R}} k(y)\,\mathrm{d}y = 1$.
(K2) There exists a $\Gamma \in (0, \infty]$ such that

$$\int_{\mathbb{R}} k(y)\mathrm{e}^{\gamma y}\,\mathrm{d}y < \infty \quad \text{for all } \gamma \in [0, \Gamma).$$

We think in particular of the cases where $k(x) = \frac{1}{2}$ for $x \in (-1, 1)$ and $= 0$ otherwise, or $k(x) = \frac{1}{2}\mathrm{e}^{-|x|}$. The last choice will lead to exact results.

To define $k_\varepsilon(x)$, pick $l_\varepsilon \to 0$ as $\varepsilon \to 0$ such that $l_\varepsilon/\varepsilon$ is an integer which tends to $\infty$ as $\varepsilon \to 0$. $k_\varepsilon(x)$ is derived from the kernel $k(\cdot)$ such that

$$k_\varepsilon(x) = \frac{\varepsilon}{l_\varepsilon} \int_{x_\varepsilon}^{x_\varepsilon + l_\varepsilon} k(y)\,\mathrm{d}y$$

for all $x \in [x_\varepsilon, x_\varepsilon + l_\varepsilon) \cap \varepsilon\mathbb{Z}$ where $x_\varepsilon \in l_\varepsilon \mathbb{Z}$. This makes $k_\varepsilon(x)$ constant over intervals of length $l_\varepsilon$. (This definition is convenient, though not necessary. But it will simplify some of our proofs.) Note that $k_\varepsilon(x)$ is the probability that an offspring whose parents are located at 0 and $\varepsilon$, is sent to $x$.

The first step in investigating the behavior for small $\varepsilon$ is to look at the limit as $\varepsilon \to 0$. The first result is called a "mean-field limit theorem" since it says that in the limit as $\varepsilon \to 0$ the density evolves as if neighboring sites are independent.

**Theorem 1.** *Suppose that $\xi_0^\varepsilon(x)$, $x \in \varepsilon\mathbb{Z}$, are independent and let $u_\varepsilon(t, x) = P(\xi_t^\varepsilon(x) = 1)$. If $u_\varepsilon(0, x) = \Phi(x)$ is a continuous function, then as $\varepsilon \to 0$, $u_\varepsilon(t, x) \to u(t, x)$ the bounded nonnegative solution of*

$$\frac{\partial u}{\partial t} = -u + \lambda(1 - u)(k * u^2) \tag{1.1}$$

*with $u(0, x) = \Phi(x)$.*

Here and in the following, $k * u^2$ denotes the convolution of $k$ and $u^2$. The theorem is not hard to prove. Analogous results for different models have been obtained by Swindle (1990), Durrett and Neuhauser (1994), and Durrett (1993). We will follow their approach in the proof of Theorem 1.

Once Theorem 1 is established, the next step is to investigate whether (1.1) admits spatially homogeneous solutions $v \in (0, 1]$ and whether or not they are stable in the sense that, for $\varepsilon$ small, the corresponding particle system has a nontrivial stationary solution whose density is close to the limiting density obtained from (1.1). On the other hand if (1.1) has only a trivial solution, i.e., $u \equiv 0$, then one expects that the corresponding particle system dies out if $\varepsilon$ is sufficiently small. The spatially homogeneous solution $v(t)$ satisfies

$$v' = -v + \lambda(1-v)v^2 \tag{1.2}$$

with spatially constant initial density $v(0)$. It is easy to see that if $\lambda < 4$, (1.2) has only the nontrivial stationary solution $v = 0$. If $\lambda > 4$ the right-hand side of (1.2) has three roots, 0, $\rho_u$ and $\rho_s$. 0 and $\rho_s$ are stable in the sense that if $0 \le v(0) < \rho_u$, then $\lim_{t\to\infty} v(t) = 0$, and if $1 \ge v(0) > \rho_u$, then $\lim_{t\to\infty} v(t) = \rho_s$. This however does not tell the whole story. Even though both 0 and $\rho_s$ are stable solutions (equilibria) of (1.2), they do not show the same behavior in the spatial setting of (1.1): one of the two stable equilibria is only "metastable." (When a metastable equilibrium is sufficiently perturbed in space, i.e., its value in a large interval is changed to the value of the stable equilibrium, the solution will settle into the stable equilibrium. The stable equilibrium however will remain unchanged under such perturbations.) The value of $\lambda$ determines which of the two equilibria is the metastable one. It turns out that there are constants $\lambda^* \ge \lambda_* \ge 4$ such that $\rho_s$ is metastable for $\lambda \in [4, \lambda_*)$ and stable for $\lambda > \lambda^*$. (We do not make any claims here about whether $\lambda_* = \lambda^*$ and about the behavior for $\lambda \in [\lambda_*, \lambda^*]$.) Therefore, the behavior of the solution is more complex when we start from a spatially inhomogeneous density $\Phi(x)$. This can already be seen when $\Phi(x)$ has $\lim_{x\to\infty} \Phi(x) = 0$ and $\lim_{x\to-\infty} \Phi(x) = \rho_s$. With this initial density, (1.1) exhibits traveling wave solutions. (This was shown by Weinberger (1982) who investigated a class of models which exhibits this behavior. Eq. (1.1) fits into this class.) Traveling waves are functions of the form $u(t,x) = U(x - c(\lambda)t)$ for some $c(\lambda)$, the wave speed corresponding to $\lambda$, with the limits $U(\pm\infty)$ existing and unequal. The wave moves to the right or to the left depending on the sign of the wave speed $c(\lambda)$. The sign of the wave speed therefore determines whether the solution converges to 0 or to $\rho_s$. This behavior carries over to the particle system. One of the main tasks will therefore be to study (1.1). We make the convention that our traveling waves $U$ always have $U(-\infty) = \rho_s$ and $U(+\infty) = 0$. This makes the wave speed $c(\lambda)$ positive (i.e., the wave moves to the right) when survival occurs. We can now characterize $\lambda_*$ and $\lambda^*$:

$$\lambda_* = \sup\{\lambda: c(\lambda) < 0\} \quad \text{and} \quad \lambda^* = \inf\{\lambda: c(\lambda) > 0\}.$$

Unfortunately, for almost all kernels $k(\cdot)$ it is impossible to say anything about numerical values for $\lambda_*$ and $\lambda^*$ or even to decide whether $\lambda_* = \lambda^*$. However, there is one exception. If we choose $k(x) = \frac{1}{2}e^{-|x|}$, (1.1) reduces to a second-order ordinary differential equation which can be integrated. We learnt this trick from Baillon who together with Théra studied a related equation (see Baillon and Théra, 1990). This is the content of the following result.

**Theorem 2.** *If* $k(x) = \frac{1}{2}e^{-|x|}$, *then* $\lambda_* = \lambda^* \equiv \lambda_0$ *and numerical estimates show that*

$$\lambda_0 \in (4.2051, 4.2052).$$

Note that $\lambda_0$ is strictly bigger than 4, that is, there is a nonempty interval of values for $\lambda$ where the system dies out but shows metastable behavior. Even though we do not know much about $\lambda_*$ and $\lambda^*$, we can find bounds on $\lambda_*$ and $\lambda^*$ by making use of a criterion in Weinberger (1982). (There is always the trivial lower bound $\lambda_* \geq 4$.) The criterion roughly says the following: If we start with a continuous function $u(0,x)$ at time 0 and if $(\partial u/\partial t)|_{t=0}$ is increasing (respectively, decreasing) everywhere, then the wave speed is positive (respectively, negative). For instance, if $k(x) = \frac{1}{2}$ for $x \in (-1, 1)$ and $= 0$ otherwise and if we let $u(0,x) = \frac{1}{2}$ for $x \leq 0$, $= \sqrt{1 - x/2}$ for $0 \leq x \leq 1$ and $= 0$ for $x \geq 1$, then the criterion yields $\lambda^* \leq \frac{16}{3}$. For the same kernel $k(x)$, i.e. $k(x) = \frac{1}{2}$ for $x \in (-1, 1)$ and $= 0$ otherwise, if we let $u(0,x) = \frac{1}{2} + \delta$ for $x < 0$, $= \delta$ for $x > N$ and linear in between, then the criterion yields that for small $\delta$ and $N$ sufficiently large, $\lambda_*$ is bounded away from 4. This last result will be useful in connection with Theorem 5.

Regardless of whether $\lambda_* = \lambda^*$ and what their values are, we can investigate the particle system and show the following theorem.

**Theorem 3.** *Suppose* $c(\lambda) > 0$ *and* $\varepsilon$ *is small. Then the particle system has a nontrivial stationary distribution in which the density of* 1*'s is close to* $\rho_s$.

In the case of negative wave speed we obtain the following result.

**Theorem 4.** *Suppose* $c(\lambda) < 0$ *and* $\varepsilon$ *is small. Then the particle system dies out locally, that is, starting from any initial configuration, the limiting distribution is the pointmass at the empty set.*

The proofs of Theorems 3 and 4 are based on a rescaling technique which was developed by Bramson and Durrett and which is reviewed in Durrett (1991). We will show that for small $\varepsilon$, the particle system will, with probability close to 1, almost do the same as the limiting deterministic system in certain finite space-time boxes when starting from certain good configurations. Iterating this will produce the desired results. Theorems 3 and 4 imply that as $\varepsilon \to 0$, $\lambda_0 \to \lambda_c$ in cases where $\lambda_* = \lambda^* \equiv \lambda_0$. In general, we believe that under the assumptions (K1) and (K2) on the kernel $k(\cdot)$, $\lambda_* = \lambda^* \equiv \lambda_0$. That is, there is only one value of $\lambda$ where the wave speed is 0. Our belief is based on numerically investigating (1.1). We do not make any conjectures about how the value of $\lambda_0$ depends on the choice of $k(\cdot)$. It is not clear whether the process survives at the critical value $\lambda_c$. However, since $\rho_s \geq \frac{1}{2}$ for $\lambda \geq 4$, that is, the density of the nontrivial equilibrium in the deterministic system is never close to 0, the phase transition in the particle system has to be very rapid when $\varepsilon$ is small.

We will now state the result which establishes metastable behavior for certain parameter values. Even though the naive mean field description suggests survival for $\lambda \in (4, \lambda_*)$, Theorem 4 shows that the process dies out. The particle system exhibits metastable behavior in this interval. This is the content of the following theorem.

**Theorem 5.** *Let $\lambda > 4$ and $\delta > 0$. If $\varepsilon$ is sufficiently small, then when starting from all sites occupied, the probability that there are less than $\varepsilon^{-1}(\rho_s - \delta)$ particles in the set $(0,1) \cap \varepsilon\mathbb{Z}$ at time $T = C' e^{\gamma'/\varepsilon}$ is smaller than $C e^{-\gamma/\varepsilon}$ where $C, C', \gamma, \gamma' \in (0, \infty)$.*

In other words, with probability close to 1 we can keep the empirical density in the unit interval (0, 1) close to $\rho_s$ for an exponentially in $\varepsilon^{-1}$ growing amount of time. That is, when looking at a specific finite interval, we have to wait at least an amount of time of order $e^{\gamma/\varepsilon}$ until instabilities set in, which drive the particle system into the trivial equilibrium. Note that Theorem 2 and the discussion after Theorem 2 showed that if $k(x) = \frac{1}{2} e^{-|x|}$ or $k(x) = \frac{1}{2}$ for $x \in (-1, 1)$ and $= 0$ otherwise, $\lambda_*$ is bounded away from 4, that is, for these choices of the kernel we showed that the interval $(4, \lambda_*)$ is not empty and hence there are kernels $k(x)$ and values of $\lambda$ where the system shows true metastable behavior.

In the proof of Theorem 5, we will make use of the fact that the birth rate is higher than the death rate in intervals where the density is between $\rho_u$ and $\rho_s$. That is, when the density in a fixed interval falls just below $\rho_s$, it will with high probability increase again.

The paper is organized as follows: In Section 2 we will construct the process and prove Theorem 1. Section 3 discusses Eq. (1.1) and proves Theorem 2. In Section 4 we will show the existence of a nontrivial equilibrium when the wave speed is positive. (This is the content of Theorem 3.) Theorem 4 will be proved in Section 5 and the metastable behavior will be investigated in Section 6. We denote constants whose values are of no interest, by $C, \gamma$ or alike. These values may change from line to line which will then be clear from the context.

## 2. Proof of Theorem 1

We will first construct the process and then prove Theorem 1. In the proof of the theorem we will adapt the arguments of Durrett (1993), and Durrett and Neuhauser (1994) to our situation. Here, a long range process is considered, whereas in the papers mentioned, processes with fast stirring are studied. This makes the arguments here somewhat different.

### 2.1. The dual process

The first step is to construct the process. For this, we introduce a number of Poisson processes, all of which are assumed to be independent. For each $x \in \varepsilon\mathbb{Z}$, let $\{U_n^x, n \geq 1\}$

be Poisson processes with rate 1. For each $x, y \in \varepsilon\mathbb{Z}$ with $k_\varepsilon(x-y) > 0$ let $\{S_n^{\varepsilon,x,y}, n \geq 1\}$ be Poisson processes with rate $\lambda k_\varepsilon(x-y)$. They have the following intrepretation: (i) at times $U_n^x$ we kill the particle at $x$ if it is present; at times $S_n^{\varepsilon,x,y}$, $y$ becomes occupied (if it is not already) if $x$ and $x+\varepsilon$ are occupied. An idea of Harris (1972) allows us to construct the process from any initial configuration. This is the standard graphical representation.

The next step is to define the dual process. We will first give an informal description of the dual process which is described in Noble's thesis (1989, 1992) and is based on an idea of Gray (1986). The dual process allows us to compute the state of a fixed site $x$ at time $t$ from the configuration at some earlier time $s$ by following the ancestral line of $x$ backwards in time starting at $(x, t)$. The dual process consists of a finite collection of finite subsets of $\varepsilon\mathbb{Z}$. The sites in those subsets are possible ancestors of $(x, t)$. Thus the computation of the state at some fixed site $(x, t)$ involves only a finite number of sites at some earlier time. We denote the dual process starting at $(x, t)$ by $\{\hat{\xi}_s^{(x,t)}\}_{s\geq 0}$ with $\hat{\xi}_0^{(x,t)} = \{x\}$. (We drop the dependence on $\varepsilon$ to avoid additional superscripts.) Strictly speaking, the dual process is only defined for $0 \leq s \leq t$. But it is easy to define it for all $s \geq 0$ by extending the graphical representation to negative times. The dual evolves as follows: If a death occurs at a point $y$ contained in some subset of $\hat{\xi}_s^{(x,t)}$ at time $t-s$, then $\hat{\xi}_s^{(x,t)}$ will be obtained from $\hat{\xi}_{s-}^{(x,t)}$ by removing any set $B$ containing $y$ from $\hat{\xi}_{s-}^{(x,t)}$. If a pair of particles located at adjacent sites $z_1$ and $z_2$ gives birth to an offspring at time $t-s$ which is then sent to a site $y$, then $\hat{\xi}_s^{(x,t)}$ will be obtained from $\hat{\xi}_{s-}^{(x,t)}$ in the following way: for each set $B \in \hat{\xi}_{s-}^{(x,t)}$ which contains $y$, add a set $B'$ which is obtained from $B$ by removing $y$ and adding $z_1$ and $z_2$. Note that we do not remove the set $B$ from the dual when adding $B'$. It follows from the construction of the dual process that the site $x$ will be occupied at time $t$ if at least one of the subsets in the dual process $\hat{\xi}_s^{(x,t)}$ is completely occupied at time $t-s$. This is the content of the following duality equation

$$x \in \xi_t^\varepsilon \iff B \subset \xi_{t-s}^\varepsilon \quad \text{for some } B \in \hat{\xi}_s^{(x,t)}. \tag{2.1}$$

As $\varepsilon \to 0$, $\hat{\xi}_s^{(x,t)}$ approaches a set valued branching process $W_s^x$. At rate 1 we remove all sets that contain the point where the death occurred. At rate $\lambda$ we add sets which are obtained from sets that contain the point where the birth occurred by removing the birth place and adding the two parents.

We will now give a precise construction of the dual process which allows us to obtain estimates on the rate of convergence to the limiting system. These are needed in the proof of Theorem 4. We will define a series of random variables from which we can compute the dual process. The first step is to define the *influence set* $\{I_\varepsilon^{x,t}(s)\}_{s\geq 0}$. Roughly speaking, $\{I_\varepsilon^{x,t}(s)\}_{s\geq 0}$ contains all sites which may influence the state of the site $(x, t)$. Here and in the following we adopt Durrett's (1993) notation and terminology. Let $I_\varepsilon^{x,t}(0) = \{x\}$. If $y \in I_\varepsilon^{x,t}(s)$ and $S_n^{\varepsilon,z,y} = t-s$ for some $n$ then we add $z$ and $z+\varepsilon$ to $I_\varepsilon^{x,t}(s)$. The arrivals $U_n^x$ have no effect on its evolution. We will now label the influence set. Let $X_\varepsilon^1(0) = x$ and $R_\varepsilon^1$ be the first time the particle at $x$ encounters an $S$ or $U$ arrival. If it encounters a $U$ arrival we set $X_\varepsilon^1 = \Delta$ to indicate that the particle died. ($\Delta$ serves as a cemetery.) In that case, the dual dies out and no further definitions are necessary. If it encounters an $S$ arrival, that is if for some $n$, $S_n^{\varepsilon,z,x} = t - R_\varepsilon^1$, then we set $X_\varepsilon^2(R_\varepsilon^1) = z$ and $X_\varepsilon^3(R_\varepsilon^1) = z + \varepsilon$. To keep track of the ancestors we define random variables $\{\mu_\varepsilon^k\}_{k\geq 1}$ and let $\mu_\varepsilon^2 = \mu_\varepsilon^3 = 1$ to indicate that 1 is the parent of 2 and 3. (Note, a parent in the influence set corresponds to an offspring in the original process.) We set $\mu_\varepsilon^1 = 0$.

By induction we can now define the process for later times. Suppose that we have defined the process up to time $R_\varepsilon^m$, $m \geq 1$. Let $K(m)$ be the total number of pairs of particles that have been created by time $R_\varepsilon^m$ and let $N_m$ be the number of particles that are alive at that time. If $N_m = 0$ there is nothing further to do. If $N_m > 0$ we simply wait until time $R_\varepsilon^{m+1}$, the first time $s > R_\varepsilon^m$ that an $S$ or $U$ arrival occurs at the location of one of the $N_m$ particles. If an $U$ arrival occurs that is, if for some $k \leq 2K(m) + 1$ and some $n$, $U_n^y = t - R_\varepsilon^{m+1}$ with $y = X_\varepsilon^k(R_\varepsilon^m)$, then we set $X_\varepsilon^k(R_\varepsilon^{m+1}) = \Delta$ and $N_{m+1} = N_m - 1$. If for some $k \leq 2K(m) + 1$ and some $n$, $S_n^{\varepsilon,z,y} = t - R_\varepsilon^{m+1}$ with $y = X_\varepsilon^k(R_\varepsilon^m)$ and $z, z + \varepsilon \notin \{X_\varepsilon^l(R_\varepsilon^m)\colon 1 \leq l \leq 2K(m) + 1\}$ then $\mu_\varepsilon^{2K(m)+2} = \mu_\varepsilon^{2K(m)+3} = k$ and $X_\varepsilon^{2K(m)+2} = X_\varepsilon^{2K(m)+3} - \varepsilon = z$, $N_{m+1} = N_m + 2$ and $K(m+1) = K(m) + 1$. If $z$ or $z + \varepsilon = X_\varepsilon^l(R_\varepsilon^{m+1})$ for some $1 \leq l \leq 2K(m) + 1$ then we say a *collision* has occurred. We will later see that collisions can be ignored: but to have everything defined we let $X_\varepsilon^{2K(m)+2}(R_\varepsilon^{m+1})$ and $X_\varepsilon^{2K(m)+3}(R_\varepsilon^{m+1})$ be two adjacent sites that are different from all the $X_\varepsilon^l(R_\varepsilon^m)$ with $1 \leq l \leq 2K(m) + 1$ if a collision occurs. Since $\bigcup_{B\in\xi_s^{(x,t)}} B \subset I_\varepsilon^{x,t}(s)$, it is clear that we can compute the state of $x$ at time $t$ from knowing the values of $\xi_{t-s}^\varepsilon(y)$ for $y \in I_\varepsilon^{x,t}(s)$ and from knowing $\{X_\varepsilon^l(s)\colon l \geq 1\}$ and $\{\mu_\varepsilon^l\colon l \geq 1\}$. We assign the value $\Upsilon$ to the location of particles which are not born yet. Note that when $X_\varepsilon^l = \Delta$ or $\Upsilon$ it cannot give birth or die.

## 2.2. *Convergence to $u(t,x)$*

Our first result is that our labeled influence set converges weakly to a labeled branching random walk:

$$(\{X_\varepsilon^k(s),\ s \geq 0\},\ \mu_\varepsilon^k)_{k\geq 1} \ \Rightarrow\ (\{X_0^k(s),\ s \geq 0\},\ \mu_0^k)_{k\geq 1} \tag{2.2}$$

where $X_0^k(s)$ is the location of the $k$th particle at time $s$ in a branching random walk on $\mathbb{R}$ where particles die at rate 1, and pairs of particles are born at rate $\lambda$. If the pair's parent is located at $x$, then the pair is sent to a site $y$ chosen at random according to the kernel $k(x - y)$. (Note that both particles are sent to the same site.) $\{X_0^k(s)\colon s \geq 0\}$ is almost described in Durrett (1979). The only difference is that here, a parent produces two instead of just one offspring.

The statement in (2.2) follows from a coupling argument. Let $X_\varepsilon \equiv \{X_\varepsilon^k(s)\colon s \geq 0\}_{k\geq 1}$ and $X_0 \equiv \{X_0^k(s)\colon s \geq 0\}_{k\geq 1}$. We couple $X_\varepsilon$ and $X_0$ in the following way: for each $k \geq 1$, $X_0^k$ uses the same exponential clocks as $X_\varepsilon^k$. That is, if the particle $X_\varepsilon^k$ dies, so does $X_0^k$; if $X_\varepsilon^k$ gives birth to a pair of offspring, so does $X_0^k$. The coupled particles $(X_\varepsilon^k, X_0^k)$ may only differ in their locations. There are two sources for this difference: one comes from collisions in the process $X_\varepsilon$, the other one comes from the fact that $X_0$

lives on $\mathbb{R}$, whereas $X_\varepsilon$ on $\varepsilon\mathbb{Z}$. The first source is easy to be dealt with since for fixed $T$, the probability of any collisions by time $T$ is negligible. To see this, let $\mathscr{X}_t = 2K(m) + 1$ for $t \in [R_\varepsilon^m, R_\varepsilon^{m+1})$. If we fix $T$, then by comparison with a branching process which produces pairs of offspring at rate $\lambda$ and in which deaths occur at rate 0, we can conclude that

$$E\mathscr{X}_t \le e^{2\lambda T}$$

for all $t \le T$. Markov inequality implies that

$$P(\mathscr{X}_T > K) \le K^{-1} e^{2\lambda T}. \tag{2.3}$$

Let $x \in \varepsilon\mathbb{Z}$ and $P$ (offspring is sent to $x$ when parents are located at 0 and $\varepsilon$) $= k_\varepsilon(x)$. Let $\kappa = \sup_{y \in \mathbb{R}} k(y)$. Then $k_\varepsilon(x) \le \kappa\varepsilon$ for all $x \in \varepsilon\mathbb{Z}$. If the total number of births by time $T$ is at most $K$, then

$$P\,(\text{collisions by time } T \mid \mathscr{X}_T \le K) \le K^2 \varepsilon\kappa. \tag{2.4}$$

If we choose $K = \varepsilon^{-0.2}$ and $T \le (0.1/2\lambda)\log 1/\varepsilon$, then the bounds in (2.3) and (2.4) tend to 0 as $\varepsilon \to 0$.

To deal with the second source of difference, we will show that on the set where no collisions occur, the locations of corresponding particles will be close. When $X_\varepsilon^k$ and $X_0^k$ give birth, they use the same random number generator to compute the location of their offspring. Let $X_\varepsilon^{l+1}$, $X_\varepsilon^{l+2}$ and $X_0^{l+1}$, $X_0^{l+2}$ be the offspring of $X_\varepsilon^k$ and $X_0^k$. Since $k_\varepsilon(\cdot)$ is constant over intervals of length $l_\varepsilon$, $|X_\varepsilon^{l+1} - X_0^{l+1}|$ and $|X_\varepsilon^{l+2} - X_0^{l+2}|$ are at most $|X_\varepsilon^k - X_0^k| + 2l_\varepsilon$. Therefore, if the number of particles born in the two processes by time $T$ is bounded by $K$, then for all $1 \le k \le K$,

$$\max_{0 \le s \le T} |X_\varepsilon^k(s) - X_0^k(s)| < 2Kl_\varepsilon.$$

If $l_\varepsilon < \varepsilon^{0.4}$, $K = \varepsilon^{-0.2}$, and $T \le (0.1/2\lambda)\log 1/\varepsilon$ as before, then $2Kl_\varepsilon \to 0$ as $\varepsilon \to 0$. This together with (2.3) and (2.4) shows that with high probability the two processes are close. Since the labels $\{\mu_\varepsilon^k\}_{k \ge 1}$ do not depend on $\varepsilon$, (2.2) follows.

We can now use the same reasoning as in Durrett (1993) to conclude that $u_\varepsilon(t,x)$ converges to a limit. We will compute $u(t,x)$ be defining a branching random walk $Y$ and independent coin flips. We will then show that this computation yields the same result as if we had used $\{X_0^k(s)\colon s \ge 0\}_{k \ge 1}$ and the initial density $\Phi(x)$. Let $Y = \{Y_s^k\colon k \ge 1\}_{s \ge 0}$ be a branching random walk on $\mathbb{R}$ starting with a single particle at $x$. $Y_s^k$ denotes the location of the $k$th particle at time $s$. We set $Y_0^1 = x$. Particles in $Y$ die at rate 1 and give birth at rate $\lambda$ to two offspring which will be sent to the same site. If a birth of two offspring occurs at time $s$ and if the total number of particles born at time $s$ is $k$, then the two offspring are called particle $k+1$ and $k+2$, respectively. Their location is at $Y_s^{k+1} = Y_s^{k+2}$. We say that a collision has occurred if offspring are sent to a site where particles of the branching process are already located. Even though, offspring of the same parent are sent to the same site, we do not call this a collision as long as the particles land on a site which was not previously chosen by the branching process. Since the process lives on $\mathbb{R}$, we can therefore say that with probability one no collisions occur in this process. If particle $k$ in the branching random walk is dead or not born yet at time $t$ we set $\zeta_0(k) = 0$. If particle $k$ is alive at time $t$ we flip a coin with probability $\Phi(Y_t^k)$ of heads and set $\zeta_0(k) = 1$ if heads come up, and $= 0$ otherwise. Using the same duality relation as for the sexual reproduction process we can compute $\zeta_t(1)$, the state of the particle at $x$ at time $t$, by working upwards in the resulting tree. We let $u(t,x) = P(\zeta_t(1) = 1)$. When there are no collisions in the branching random walk $\{X_\varepsilon^k(s): s \geq 0\}_{k \geq 1}$, then the family structure of the influence set $\{I_\varepsilon^{x,t}(s): s \geq 0\}$ and of the branching random walk $\{Y_s\}_{s \geq 0}$ are the same. Furthermore, the computation of the states of $\zeta_t(1)$ and $(x,t)$ in the two processes yields the same result in the limit. Since we assumed that the initial density is continuous, it follows from the continuous mapping theorem that $u_\varepsilon(t,x) \to u(t,x)$. It is straightforward to conclude the following result.

**Lemma 2.1** *If $x_\varepsilon \to x$, then $u_\varepsilon(t,x_\varepsilon) \to u(t,x)$.*

This shows that $u_\varepsilon(t,x)$ converges to $u(t,x)$ uniformly on compact sets.

The proof of the last result also implies that dual processes for different sites are asymptotically independent. That is, if $x_\varepsilon \to x$ and $y_\varepsilon \to y$,

$$P(x_\varepsilon \notin \xi_t^\varepsilon;\ y_\varepsilon, y_\varepsilon + \varepsilon \in \xi_t^\varepsilon) \to (1 - u(t,x))u^2(t,y), \quad \text{as } \varepsilon \to 0. \tag{2.5}$$

*2.3. The limit satisfies the integro-differential equation*

The following argument is adapted from Swindle (1990). We need to define $u_\varepsilon(x,t)$ for all $x \in \mathbb{R}$. It is already defined for $x \in \varepsilon\mathbb{Z}$. If $x_\varepsilon \in \varepsilon\mathbb{Z}$, and $x \in [x_\varepsilon, x_\varepsilon + \varepsilon)$, then define $u_\varepsilon(x,t) = u_\varepsilon(x_\varepsilon, t)$.

Let $x_\varepsilon \in \varepsilon\mathbb{Z}$. A simple generator calculation yields:

$$\frac{\partial u_\varepsilon(t,x_\varepsilon)}{\partial t} = -u_\varepsilon(t,x_\varepsilon) + \lambda \sum_{y \in \varepsilon\mathbb{Z}} k_\varepsilon(x_\varepsilon - y) P(x_\varepsilon \notin \xi_t^\varepsilon;\ y, y + \varepsilon \in \xi_t^\varepsilon). \tag{2.6}$$

Since the summands on the left-hand side are bounded and according to (2.5) converge pointwise to the corresponding integrand on the right-hand side, it follows that

$$\lambda \sum_{y \in \varepsilon\mathbb{Z}} k_\varepsilon(x_\varepsilon - y) P(x_\varepsilon \notin \xi_t^\varepsilon;\ y, y + \varepsilon \in \xi_t^\varepsilon) \to \lambda(1 - u(t,x)) \int k(x-y) u^2(t,y)\,\mathrm{d}y.$$

To see that

$$\frac{\partial u_\varepsilon(t,x_\varepsilon)}{\partial t} \to \frac{\partial u(t,x)}{\partial t}, \tag{2.7}$$

we write the left-hand side of (2.7) in integral form,

$$u_\varepsilon(t,x_\varepsilon) = u_\varepsilon(0,x_\varepsilon) + \int_0^t \frac{\partial u_\varepsilon(s,x_\varepsilon)}{\partial s}\,\mathrm{d}s \tag{2.8}$$

and show that for any given $x_\varepsilon$, $\partial u_\varepsilon(s,x_\varepsilon)/\partial s$ converges uniformly on $s\in[0,t]$ as $\varepsilon\to 0$. (This is basically (3.6) in Swindle (1990).) (2.7) then follows from the fact that $u_\varepsilon(0,x_\varepsilon)\to u(0,x)$ as $\varepsilon\to 0$.

To show uniform convergence of $\partial u_\varepsilon(s,x_\varepsilon)/\partial s$ on $s\in[0,t]$, let $u_1\equiv u_\varepsilon(s,x_\varepsilon)$ and $\sum^{(1)}=\sum_{y\in\varepsilon\mathbb{Z}} k_\varepsilon(x_\varepsilon-y)u_\varepsilon^2(s,y)$. ($u_2$ and $\sum^{(2)}$ are the corresponding quantities when $\varepsilon$ is replaced by $\varepsilon'$.) Then

$$\begin{aligned}
\frac{\partial u_1}{\partial s}-\frac{\partial u_2}{\partial s} &= \lambda\sum_{y\in\varepsilon\mathbb{Z}} k_\varepsilon(x_\varepsilon-y)[P(x_\varepsilon\notin\xi_s^\varepsilon;\, y,y+\varepsilon\in\xi_t^\varepsilon)-(1-u_1)u_\varepsilon^2(s,y)]\\
&\quad-\lambda\sum_{y\in\varepsilon'\mathbb{Z}} k_{\varepsilon'}(x_{\varepsilon'}-y)[P(x_{\varepsilon'}\notin\xi_s^{\varepsilon'};\, y,y+\varepsilon'\in\xi_t^{\varepsilon'})-(1-u_1)u_{\varepsilon'}^2(s,y)]\\
&\quad+\lambda(1-u_1)\textstyle\sum^{(1)}-\lambda(1-u_2)\sum^{(2)}-(u_1-u_2)\\
&= \lambda\sum_{y\in\varepsilon\mathbb{Z}} k_\varepsilon(x_\varepsilon-y)[P(x_\varepsilon\notin\xi_s^\varepsilon;\, y,y+\varepsilon\in\xi_t^\varepsilon)-(1-u_1)u_\varepsilon^2(s,y)]\\
&\quad-\lambda\sum_{y\in\varepsilon'\mathbb{Z}} k_{\varepsilon'}(x'_\varepsilon-y)[P(x_{\varepsilon'}\notin\xi_s^{\varepsilon'};\, y,y+\varepsilon'\in\xi_t^{\varepsilon'})-(1-u_1)u_{\varepsilon'}^2(s,y)]\\
&\quad+\lambda(u_2-u_1)\textstyle\sum^{(1)}+\lambda(1-u_2)\left(\sum^{(1)}-\sum^{(2)}\right)+(u_2-u_1).
\end{aligned}$$

Since $|\sum^{(1)}|\le 1$, an application of the triangle inequality, Lemma 2.1, and (2.5) shows that the first four terms tend to 0 as $\varepsilon\to 0$ uniformly for all $s\in[0,t]$. For the difference $\sum^{(1)}-\sum^{(2)}$, observe that $|u_2|\le 1$ and that

$$\textstyle\sum^{(1)}-\sum^{(2)}=\left(\sum^{(1)}-k*u_1^2\right)+\left(k*u_2^2-\sum^{(2)}\right)+(k*(u_1^2-u^2))+(k*(u^2-u_2^2)).$$

The first two terms on the right-hand side vanish in the limit as $\varepsilon\to 0$. The dominated convergence theorem implies that the last two terms also vanish in the limit. Since all four terms converge to 0 uniformly on $s\in[0,t]$, uniform convergence is established and the proof of Theorem 1 is finished.

## 3. Proof of Theorem 2

In this section we study Eq. (1.1). Results on its qualitative behavior are based on Weinberger (1982). Then, we concentrate on a special choice of $k(\cdot)$ and discuss the behavior of the system for this kernel. This will constitute the proof of Theorem 2.

### 3.1. Weinberger's results

Weinberger (1982) discusses a class of population models to which our model belongs. Results of his paper imply:
(i) Our system exhibits traveling wave solutions.
(ii) A shape theorem holds.
We will first explain (i) and (ii). Recall Eq. (1.1):

$$\frac{\partial u}{\partial t} = -u + \lambda(1-u)(k * u^2),$$

with initial condition $u(0,x) = \Phi(x)$. To explain (i), let $\Phi(x)$ be continuous and nonincreasing with $\Phi(-\infty) \in (\rho_u, 1)$, $\Phi(x) > 0$ for $x < 0$, and $\Phi(x) = 0$ for $x \geq 0$. Then, for fixed $\lambda$, there is a real number $c = c(\lambda)$ so that $u(t, \gamma t) \to 0$ for $\gamma > c$ and $u(t, \gamma t) \to \rho_s$ for $\gamma < c$. $c(\lambda)$ is called wave speed. To explain (ii), start from an initial density $\Phi(x)$ which has the following properties: $\Phi(x) \in (\rho_u, 1]$ for $x \in C$, a compact set centered at the origin, and $\Phi(x) = 0$ outside a bounded set that contains $C$. Then, if $c(\lambda) > 0$,

$$\lim_{t \to \infty} \min_{x : |x| \leq (1-\varepsilon)tc(\lambda)} u(t,x) = \rho_s. \tag{3.1}$$

This says that the disturbance at time 0 expands at a constant speed and fills most of the set $(-tc(\lambda), tc(\lambda))$. Attractiveness implies a bit more about the wave speed $c(\lambda)$.

**Lemma 3.1.** *$c(\lambda)$ is nondecreasing in $\lambda$.*

**Proof.** Let $u_1$ correspond to the system with $\lambda_1$ and $u_2$ to the one with $\lambda_2$ and assume $\lambda_1 \leq \lambda_2$. Then, the lemma will follow upon establishing that if $u_1(0,x) \leq u_2(0,x)$ then $u_1(t,x) \leq u_2(t,x)$. Since $\lambda_1 \leq \lambda_2$,

$$\frac{\partial u_2}{\partial t} \geq -u_2 + \lambda_1(1-u_2)(k * u_2^2), \tag{3.2}$$

which shows that $u_2$ is a supersolution, i.e., $u_2(t,x) \geq u_1(t,x)$ when starting with $u_2(0,x) \geq u_1(0,x)$. Both solutions converge to traveling wave fronts. Since $u_2(t,x) \geq u_1(t,x)$, it follows that the wave speed corresponding to $\lambda_2$ is at least as large as the one corresponding to $\lambda_1$. □

Note that Lemma 3.1 does not exclude the existence of whole intervals where the wave speed $c(\lambda)$ is constant. In particular, there might be cases where $c(\lambda) = 0$ for $\lambda \in (\lambda_*, \lambda^*)$ and $\lambda_* \neq \lambda^*$. ($\lambda_*$ and $\lambda^*$ were defined in the introduction.)

The rest of Section 3 is devoted to the proof of Theorem 2.

### 3.2. A special choice for the kernel

As already mentioned in the introduction, there is a criterion in Weinberger (1982) that allows us to find bounds on $\lambda^*$, the smallest value of $\lambda$ for which the wave speed is positive. We applied it to the kernel $k(x) = \frac{1}{2}$ for $x \in (-1, 1)$ and $= 0$ otherwise. This criterion did not give very good bounds in that case, so we will not pursue this any further. The structure of (1.1) makes it difficult to obtain good bounds on $\lambda_*$ and $\lambda^*$. However, if we let $k(x) = \frac{1}{2}e^{-|x|}$, then (1.1) reduces to an ordinary differential equation which is tractable. Its investigation shows $\lambda_* = \lambda^*$ and enables us to compute its value.

For the rest of this section, let $k(x) = \exp(-|x|)/2$. We will first discuss the stationary solution of (1.1) with that choice of $k(x)$ and compute the value of $\lambda$ where $c(\lambda) = 0$. We will then show that there is exactly one value of $\lambda$ where $c(\lambda) = 0$. The stationary solution of (1.1) satisfies

$$\lambda(1-u)(k * u^2) = u. \tag{3.3}$$

Choosing $k(x) = e^{-|x|}/2$ and differentiating (3.3) twice reduces (3.3) to

$$u''(1-u) + 2(u')^2 = u(1-u)^2 - u^2\lambda(1-u)^3. \tag{3.4}$$

This equation can be found in Kamke (1948). Multiplying (3.4) with $2(1-u)^{-5}u'$ and using the fact that

$$[(1-u)^{-4}(u')^2]' = 2u''(1-u)^{-4}u' + 4(1-u)^{-5}(u')^3 \tag{3.5}$$

allows us to integrate (3.4). Carrying out the integration gives

$$(u')^2 = (1-u)^2\{(1-u)^2[C - 4\lambda\log(1-u) - 2\lambda u] + 1 - 2(1+\lambda)(1-u)\} \tag{3.6}$$

where $C$ is the constant of integration. The first observation is that if $\lim_{x\to\pm\infty} u(x)$ exists, then $\lim_{x\to\pm\infty} u'(x) = 0$. The proof is trivial and we omit it.

To identify those values of $\lambda$ that permit such solutions, we choose $C$ in (3.6) so that $u(-\infty) = \rho_s$ and $u'(-\infty) = 0$. Since $\rho_s$ depends solely on $\lambda$, $C$ is only a function of $\lambda$ and can be easily computed. We are interested in those values of $\lambda$ for which $u'(+\infty) = u(+\infty) = 0$. Computing $(u')^2$ in (3.6) for our choice of $C$ and $u = 0$ for different values of $\lambda$ shows that there exists a value $\lambda_0$ so that if $\lambda < \lambda_0$, $(u')^2$ is negative, if $\lambda = \lambda_0$, $(u')^2 = 0$, and if $\lambda > \lambda_0$, $(u')^2 > 0$. This implies that there is only one value of $\lambda$, namely $\lambda_0$, for which $u(-\infty) = \rho_s$, $u(+\infty) = 0$ and $u'(\pm\infty) = 0$. This is the traveling wave with speed 0 we are seeking. The value of $\lambda_0$ can be computed from this and one finds

$$\lambda_0 \in (4.2051, 4.2052). \tag{3.7}$$

This and Lemma 3.1 thus show that there is only one value of $\lambda$ where $c(\lambda) = 0$. This proves Theorem 2.

## 4. Proof of Theorem 3

In this section we will show the existence of a nontrivial stationary distribution when the wave speed is positive. The proof is based on a rescaling argument. This technique was developed by Bramson and Durrett and is reviewed in Durrett (1991). It is by now a standard technique and has been frequently applied (see e.g., Bramson, 1989; Bramson and Durrett, 1988). The heart of the argument is to show that if $0 < p < 1$ and $\varepsilon$ is small, the process when viewed on suitable time and length scales, dominates a weakly dependent oriented site percolation process where sites are wet with probability $p$. The comparison will be done in three steps. First, we will use Weinberger's (1982) shape theorem stated in the last section. It implies that if initially the density in a sufficiently large interval exceeds $\rho_u$, then $T$ units of time later ($T$ sufficiently large), there will be a larger interval (containing the original interval) in which the density will be close to $\rho_s$. In the second step, we compare the particle system with the limiting system. The mean field theorem and a second moment computation will show that with probability close to 1, our particle system will do almost the same as the limiting system. Iterating this and comparing with a mildly dependent oriented site percolation process will finally produce the desired result.

Let $\delta < (\rho_s - \rho_u)/10$. The first step is to observe that (3.1) implies the following lemma.

**Lemma 4.1.** *We can pick* $T < \infty$ *so that if* $u(0,x) \geq \rho_u + \delta$ *on* $[-L, L]$ *then* $u(T,x) \geq \rho_s - \delta$ *on* $[-4L, 4L]$.

Lemma 4.1 shows that if we start the deterministic system with density at least $\rho_u + \delta$ in $[-L, L]$ then at time $T$ we will have density at least $\rho_s - \delta$ in $2L + [-L, L]$ and $-2L + [-L, L]$.

In the next step we will use Theorem 1 supplemented by a second moment computation to show that with probability close to 1, the particle system exhibits a similar behavior. To carry this out, we divide $[-L, L]$ into small subintervals whose lengths go to 0 as $\varepsilon \to 0$. We choose the length so that at the same time the number of sites in each subinterval goes to infinity. If we start with "sufficiently" many particles in all of the subintervals in $[-L, L]$, then with high probability, we will obtain "enough" particles in all of the small subintervals in $-2L + [-L, L]$ and $2L + [-L, L]$ by time $T$. To make this precise we need to define an empirical density for the particle system. For this let $J = [-L, L]$ and divide $J$ into $\lfloor 2L/l_\varepsilon \rfloor$ subintervals of length $l_\varepsilon$. $\lfloor \cdot \rfloor$ denotes the integer part. For $x \in l_\varepsilon \mathbb{Z}$, let

$$\begin{aligned} \tilde{\xi}_t^{1,\varepsilon}(x) &= (\varepsilon/l_\varepsilon) \sum_{y \in x + [0, l_\varepsilon)} \xi_t^\varepsilon(y), \\ \tilde{\xi}_t^{2,\varepsilon}(x) &= (\varepsilon/l_\varepsilon) \sum_{y \in x + [0, l_\varepsilon)} \xi_t^\varepsilon(y)\, \xi_t^\varepsilon(y + \varepsilon) \end{aligned} \tag{4.1}$$

be the empirical one-point and two-point densities of particles in the interval $x + [0, l_\varepsilon)$. For simplicity we will assume that $l_\varepsilon/\varepsilon$ is an integer. Furthermore, we choose $l_\varepsilon$ so that the number of subintervals in $J$ and the number of sites in each subinterval tend to infinity. We will say more about the choice of $l_\varepsilon$ below.

To make the comparison with a mildly dependent oriented site percolation we have to make sure that the particles we use at time $T$ in the intervals $-2L + [-L, L]$ and $2L + [-L, L]$ do not have ancestors that are too far away. We will therefore modify the empirical densities $\tilde{\xi}_t^{1,\varepsilon}(x)$and $\tilde{\xi}_t^{2,\varepsilon}(x)$. We call the site $y$ *good* at time $t$ if the influence set $I_\varepsilon^{y,t}(s)$ does not escape from $y + (-NL, NL)$ for all $s \in [0, T]$. $N$ will be a large positive integer which we will specify below. The modified empirical one-point and two-point densities are then defined as

$$\begin{aligned}
\eta_t^{1,\varepsilon}(x) &= (\varepsilon/l_\varepsilon) \sum_{\substack{y \in x + [0, l_\varepsilon) \\ y \text{ good}}} \xi_t^\varepsilon(y), \\
\eta_t^{2,\varepsilon}(x) &= (\varepsilon/l_\varepsilon) \sum_{\substack{y \in x + [0, l_\varepsilon) \\ y \text{ and } y+\varepsilon \text{ are good}}} \xi_t^\varepsilon(y)\, \xi_t^\varepsilon(y+\varepsilon).
\end{aligned} \tag{4.2}$$

It is clear that for any $\delta > 0$ we can choose $N$ large enough so that the probability of a site being good is at least $1 - \delta$.

We will first show that the particle system does almost the same as the limiting system when starting from product measure. This is the content of the following lemma.

**Lemma 4.2.** *Suppose $\xi_0^\varepsilon(x)$ are independent with $P(\xi_0^\varepsilon(x) = 1) = \Phi(x)$, continuous, and $\Phi(x) \geq \rho_u + \delta$ for $x \in [-L, L]$. If $\varepsilon$ is small then with probability close to 1, $\eta_T^{1,\varepsilon}(x) \geq \rho_s - 4\delta$ for all $x \in l_\varepsilon\mathbb{Z} \cap ([-3L, -L] \cup [L, 3L]$.*

**Proof.** We will prove this by computing the mean and the variance of the empirical density in the small subintervals and then using Chebyshev's inequality. Theorem 1, Lemma 4.1 and the definition of a site being good imply that

$$E\eta_T^{1,\varepsilon}(x) = (\varepsilon/l_\varepsilon) \sum_{y \in x + [0, l_\varepsilon)} P(\xi_T^\varepsilon(y) = 1, y \text{ good}) \geq \rho_s - 3\delta \tag{4.3}$$

for $x \in l_\varepsilon\mathbb{Z} \cap [L, 3L]$ and for $\varepsilon$ sufficiently small.

It follows from a remark in Section 2 about dual processes for different sites being asymptotically independent that as $\varepsilon \to 0$

$$c_\varepsilon = \sup_{x \neq y} \operatorname{cov}(I_{\{\xi_T^\varepsilon(x) = 1\}}, I_{\{\xi_T^\varepsilon(y) = 1\}}) \to 0. \tag{4.4}$$

($I_A$ denotes the indicator function of $A$). We will now compute the variance

$$\mathrm{var}(\eta_T^{1,\varepsilon}(x)) = (\varepsilon/l_\varepsilon)^2 \Bigg[ \sum_{\substack{y \in x + [0, l_\varepsilon) \\ y \text{ good}}} \mathrm{var}(\xi_T^\varepsilon(y))$$

$$+ \sum_{\substack{y, x \in x + [0, l_\varepsilon) \\ y \neq x, \text{ both good}}} \mathrm{cov}(I_{\{\xi_T^\varepsilon(y) = 1\}}, I_{\{\xi_T^\varepsilon(z) = 1\}}) \Bigg]. \tag{4.5}$$

Since $\xi_T^\varepsilon(y) \in \{0, 1\}$, the variance of $\xi_T^\varepsilon(y)$ is not greater than $\frac{1}{4}$. This together with (4.4) yields

$$\mathrm{var}(\eta_T^{1,\varepsilon}(x)) \leq \frac{\varepsilon}{4l_\varepsilon} + c_\varepsilon. \tag{4.6}$$

An application of Chebyshev's inequality then gives

$$P(\eta_T^{1,\varepsilon}(x) < \rho_s - 4\delta \text{ for some } x \in l_\varepsilon\mathbb{Z} \cap [L, 3L]) \leq \frac{1}{\delta^2}\left(\frac{\varepsilon}{4l_\varepsilon} + c_\varepsilon\right)(2L/l_\varepsilon). \tag{4.7}$$

If we let $l_\varepsilon \to 0$ slowly enough so that

$$\varepsilon/l_\varepsilon^2 \to 0 \quad \text{and} \quad c_\varepsilon/l_\varepsilon \to 0, \tag{4.8}$$

then the right-hand side of (4.7) goes to 0 as $\varepsilon$ tends to 0. This finishes the proof of Lemma 4.2. □

When starting from product measure at time 0, the distribution of the configuration at time $T$ is not product measure. Therefore, in order to iterate the procedure, we must allow other initial configurations as well. To do this, we will now show that starting from certain configurations with sufficiently high density is almost the same as starting from product measure as in Lemma 4.2. We will call a configuration *nice* in the interval $I$ if $\eta_t^{1,\varepsilon}(x) \geq \rho_u + 3\delta$ for all $x \in l_\varepsilon\mathbb{Z} \cap I$ and if $\eta_t^{2,\varepsilon} \geq (\rho_u + 2\delta)^2$ for all $x \in l_\varepsilon\mathbb{Z} \cap I$. We will now prove Lemma 4.3.

**Lemma 4.3.** *Starting from a nice configuration in* $[-L, L]$ *at time* 0 *is not much worse than starting from product measure with density* $\geq \rho_u + \delta$. *That is, if* $\zeta_0^\varepsilon$ *denotes the process starting from product measure with density* $\rho_u + \delta$, *and* $\xi_0$ *is a nice configuration in* $[-L, L]$ *at time* 0. *Then*

$$P(\xi_T^\varepsilon(x) = 1) + \delta \geq P(\zeta_t^\varepsilon(x) = 1).$$

*Furthermore, starting with a nice configuration in* $[-L, L]$ *produces, with probability close to* 1, *nice configurations in both* $[-3L, -L]$ *and* $[L, 3L]$.

**Proof.** The definition of the kernel $k_\varepsilon(\cdot)$ implies that when choosing a pair of parents in the dual process, we basically first choose one of the subintervals of length $l_\varepsilon$ and then a pair within the subinterval. The pair within the subinterval is chosen at random with equal probabilities for all pairs. If the initial configuration is nice and if all the pairs of parents live in different subintervals, then this is at least as good as if we had started from a product measure with density $\rho_u + 2\delta$. To estimate the probability of what can go wrong, we will first bound the size of the dual process and then estimate how likely it is that parents will be sampled more than once from the same subinterval or that they will be sampled from boundaries of the subinterval. To estimate the size of the dual we will use results in Durrett and Neuhauser (1994) which imply

$$P(\mathscr{X}_T > K) \le C\,\frac{e^{4\lambda T}}{K^2}, \tag{4.9}$$

where $\mathscr{X}_T$ was defined in Section 2. On the set where we have at most $K$ parents, the probability that parents will be sampled more than once from the same subinterval or that they will be sampled from the boundary of a subinterval can be bounded by

$$\le 2K^2 l_\varepsilon \sup_x k(x). \tag{4.10}$$

By first choosing $K$ large and then $\varepsilon$ small we can make the probability of the bad events in (4.9) and (4.10) smaller than $\delta$. We will now compare the process starting from a nice configuration with a process starting from product measure with density $\rho_u + \delta$. Denote the first process by $\xi_t^\varepsilon$ and the second by $\zeta_t^\varepsilon$. From the above we see that

$$P(\xi_T^\varepsilon(x) = 1) + \delta \ge P(\zeta_T^\varepsilon(x) = 1).$$

All that is left to show is that a nice configuration in $[-L, L]$ will with high probability produce a nice configuration in both $[-3L, -L]$ and $[L, 3L]$. This will follow from estimates which are similar to the ones in the proof of Lemma 4.2. Recall the modified empirical two-point density

$$\eta_t^{2,\varepsilon}(x) = (\varepsilon/l_\varepsilon) \sum_{\substack{y \in x + [0, l_\varepsilon) \\ y \text{ and } y+\varepsilon \text{ are good}}} \xi_t^\varepsilon(y)\,\xi_t^\varepsilon(y+\varepsilon)$$

for $x \in l_\varepsilon \mathbb{Z}$. Then, as before

$$E\eta_T^{2,\varepsilon}(x) \ge (\rho_s - 4\delta)^2$$

and

$$\begin{aligned}\operatorname{var}(\eta_T^{2,\varepsilon}(x)) = (\varepsilon/l_\varepsilon)^2 \Bigg\{ &\sum_{\substack{y \in x + [0, l_\varepsilon) \\ y,\, y+\varepsilon \text{ are good}}} \operatorname{var}(\xi_T^\varepsilon(y)\,\xi_T^\varepsilon(y+\varepsilon)) \\ &+ \textstyle\sum^{(1)} \operatorname{cov}(\xi_T^\varepsilon(y)\xi_T^\varepsilon(y+\varepsilon), \xi_T^\varepsilon(z)\xi_T^\varepsilon(z+\varepsilon)) \\ &+ 2\textstyle\sum^{(2)} \operatorname{cov}(\xi_T^\varepsilon(y)\xi_T^\varepsilon(y+\varepsilon), \xi_T^\varepsilon(y+\varepsilon)\xi_T^\varepsilon(y+2\varepsilon)) \Bigg\}.\end{aligned}$$

The $\sum^{(1)}$ runs over all those $y$, $y + \varepsilon$, $z$, $z + \varepsilon$ which are all different and good. The $\sum^{(2)}$ takes then care of the terms where the sites overlap. If we denote by $d_\varepsilon$ the covariance of the indicators which do not overlap then as in the estimate of the one-point densities, $d_\varepsilon \to 0$ as $\varepsilon \to 0$. The reason is the same as before: The covariance goes to 0 since the duals do not interfere with each other once $\varepsilon$ is sufficiently small. Then, we can bound this by

$$\le (\varepsilon/l_\varepsilon)^2\, C(l_\varepsilon/\varepsilon) + d_\varepsilon,$$

where we combined the first and the third sum into the first term. An application of Chebyshev's inequality gives

$$P(\eta_T^{2,\varepsilon}(x) < (\rho_s - 5\delta)^2 \quad \text{for some } x \in l_\varepsilon\mathbb{Z} \cap [L, 3L])$$

goes to 0 as $\varepsilon \to 0$. The estimates for the one-point density are similar to the ones in the proof of Lemma 4.2 and we omit the details. □

With Lemma 4.3 established, Theorem 3 follows from a comparison with $2N$-dependent oriented site percolation where the probability of a site being open is at least $p$. We will now describe this percolation process. Let $\mathscr{L} = \{(x,n) \in \mathbb{Z}^2\colon x + n \text{ is even}\}$. We define random variables $\omega(x,n)$, $(x,n) \in \mathscr{L}$, to be 1 if $(x,n)$ is *open* and to be 0 if $(x,n)$ is *closed*. We say that $(y,n)$ can be reached from $(x,m)$ and write $(x,m) \to (y,n)$ if there is a sequence of points $x = x_m, \ldots, x_n = y$ so that $|x_l - x_{l+1}| = 1$ for $m \le l < n$ and $\omega(x_l, l) = 1$ for $m \le l \le n$. Let $C_0 = \{z\colon (0,0) \to z\}$ be the cluster containing $(0,0)$ and let $\Omega_\infty = \{|C_0| = \infty\}$ be the event that percolation occurs. *$2N$-dependent with density at least $p$* means that if $(x_i, n_i)$, $1 \le i \le m$, is a sequence with $|x_i - x_j| > 2N$ whenever $i \ne j$ and $n_i = n_j$, then

$$P(\omega(x_i, n_i) = 0 \quad \text{for all } 1 \le i \le m) = (1-p)^m,$$

that is, boxes that are separated by more than $2N$ at the same time level are independent. A straightforward extension of results in Section 10 of Durrett (1984) shows the following result.

**Lemma 4.4.** *If $P(\omega(x,m) = 1) \ge p$ and $p \ge 1 - 6^{-4(2N+1)^2}$ then $P(\Omega_\infty) > 0$.*

To construct a stationary distribution for the particle system we start the particle system from a product measure in which $P(\xi_t^\varepsilon(x) = 1) = \rho_s$. We take the Cesaro average of its distribution at time $0 \le s \le S$, and extract a convergent subsequence. Since our process has the Feller property the limit $\mu$ is a stationary distribution. See Proposition 1.8(d) of Liggett (1985). The definition of $\mu$ implies that it is translation invariant. To check that it is nontrivial for small $\varepsilon$ we will use the comparison with oriented percolation.

We say that $(x,n) \in \mathscr{L}$ is *occupied* if $\eta_{nT}^{1,\varepsilon}(y) \ge \rho_u + 3\delta$ and $\eta_{nT}^{2,\varepsilon}(y) \ge (\rho_u + 2\delta)^2$ for all $y \in l_\varepsilon\mathbb{Z} \cap x + [-L, L]$. Let $V_n = \{x\colon (x,n) \text{ is occupied}\}$ be the set of occupied sites at time $n$ and suppose that the set of occupied sites at time 0 in the percolation process is $W_0 = V_0$. We will define random variables $\omega(x,n)$ by induction so that if $W_n = \{y: (x,0) \to (y,n)$ for some $x \in W_0\}$ is the set of *wet* sites at time $n$ then $W_n \subset V_n$. If $x-1$ or $x+1$ is in $V_{n-1}$ we let $\omega(x,n) = 1$ if $x \in V_n$, $= 0$ otherwise. If neither $x-1$ nor $x+1$ is in $V_{n-1}$ we define $\omega(x,n)$ by flipping an independent coin with probability $p$ of heads (1) and probability $1-p$ of tails (0). It follows from the Markov property and the choice of $N$ in the definition of a site being good that if we condition on $\mathscr{F}_{(n-1)T}$, the information known at time $(n-1)T$, then $\omega(x,n)$ for a fixed value of $n$ are $2N$-dependent and take the value 1 with probability at least $p$, so it follows by induction that the whole collection has these properties.

It is easy to check by induction that $W_n \subset V_n$. Since $W_n$ is the set of wet sites at time $n$ in a supercritical percolation process that starts from an initial product measure, results in Durrett (1984) imply that

$$\inf_n P(\text{some site } (2k,2n) \text{ with } |k| \le K \text{ is wet}) \ge 1\text{-}\gamma_K, \tag{4.11}$$

with $\gamma_K \to 0$ as $K \to \infty$ (see appendix of Durrett and Neuhauser (1991)). (4.11) and the definition of occupied imply that the translation invariant stationary distribution $\mu$ that we have constructed is nontrivial and the proof of Theorem 3 is complete.

## 5. Nonexistence of nontrivial stationary distributions

In this section we will show that if the birthrate is too low, then the particle system dies out locally with probability 1. Again, we will compare the particle system with the limiting deterministic system and make use of results in Weinberger (1982). Another application of the rescaling technique introduced in Section 4 will then complete the proof.

Associated with $\lambda$ is a wave speed $c(\lambda)$. We are now interested in those values of $\lambda$ where the wave speed is negative. Let $v = -c(\lambda)/2$. Let $\Phi_L(x) = \rho_u/2$ for $x \in [-L+1, L-1]$, $= 1$ for $x \notin [-L, L]$, and smooth in between. Let $u_L(t,x)$ denote the solution of (1.1) with initial condition $\Phi_L$. It follows from Theorem 6.2 in Weinberger (1982) that there are constants $0 < \gamma_0, C_0 < \infty$ so that

$$u_L(T,x) \le C_0\,\varepsilon^{\gamma_0} \quad \text{for } x \in [-L-vT-1, L+vT+1], \tag{5.1}$$

where we choose $T = A_0 \log(1/\varepsilon)$. We will specify $A_0$ below. As in Section 4 (cf., formula (4.1)), we define the empirical densities of particles in subintervals of the form $x + [0, l_\varepsilon)$ for $x \in l_\varepsilon\mathbb{Z}$. The empirical one-point density for $x \in l_\varepsilon \mathbb{Z}$ is

$$\tilde{\xi}_t^{1,\varepsilon}(x) = (\varepsilon/l_\varepsilon) \sum_{y \in x + [0, l_\varepsilon)} \xi_t^\varepsilon(y). \tag{5.2}$$

The empirical two-point density for $x \in l_\varepsilon \mathbb{Z}$ is

$$\tilde{\xi}_t^{2,\varepsilon}(x) = (\varepsilon/l_\varepsilon) \sum_{y \in x + [0, l_\varepsilon)} \xi_t^\varepsilon(y)\,\xi_t^\varepsilon(y+\varepsilon). \tag{5.3}$$

We say that $\xi_t^\varepsilon$ has empirical one-point (resp. two-point) density $\rho$ (resp. $\rho^2$) in $I$ if this holds for all subintervals $x + [0, l_\varepsilon)$, $x \in I \cap l_\varepsilon \mathbb{Z}$. We will prove Lemma 5.1.

**Lemma 5.1.** *There are constants $0 < \gamma_1, C_1, L_0 < \infty$ so that if $\xi_0^\varepsilon$ is a configuration which has empirical one-point density at most $\rho_u/4$ and empirical two-point density at most $(\rho_u/3)^2$ in $[-L, L]$, $L \geq L_0$, then*

$$P(\xi_T^\varepsilon(x) = 1) \leq C_1 \varepsilon^{\gamma_1} \quad \text{for } x \in [-L - vT, L + vT].$$

**Proof.** We will show that starting from product measure with density $\rho_u/3$ in $[-L, L]$ is not much worse than starting from configurations described in the lemma. Let $\zeta_t^\varepsilon$ denote the configuration at time $t$ when starting from a product measure with density $\rho_u/3$ in $[-L, L]$ and density 1 outside of $[-L, L]$. Because of attractiveness, this will dominate all other product measures with density $\rho_u/3$ in $[-L, L]$ and density at most 1 outside of $[-L, L]$. The dual process $I_\varepsilon^{x,T}(t)$, $t \leq T$, is close to a limiting branching random walk. Let $\mathscr{X}_t$ be the number of particles in a branching process at time $t$ that produces two offspring simultaneously at rate $\lambda$. Since $E\mathscr{X}_t \leq 2\mathrm{e}^{2\lambda t}$, it follows from the choice of $T$ that

$$P(\mathscr{X}_T \geq \varepsilon^{-2\alpha}) \leq 2\varepsilon^{2\alpha - 2\lambda A_0} \leq \varepsilon^{\alpha} \tag{5.4}$$

if $\alpha > 2\lambda A_0$. If the number of particles in the process is less than $\varepsilon^{-2\alpha}$, then the probability of a collision is small. Using (2.4), we can bound this probability by

$$\leq \varepsilon^{-2\alpha} \varepsilon^{-2\alpha} \varepsilon \kappa \to 0 \tag{5.5}$$

if $1 - 4\alpha > 0$. When no collisions occur, the structure of the dual process is the same as that of the branching random walk. If we have less than $\varepsilon^{-2\alpha}$ births in $\mathscr{X}_T$ by time $T$, it follows from the discussion after (2.4) that all the particles in the dual process are within $2\varepsilon^{-2\alpha} l_\varepsilon$ of their counterparts in the limiting process. We then assign values to each particle in the dual process and its counterpart in the limiting process as in Section 2. The two values can only differ if the particle in the limiting process lands within $2\varepsilon^{-2\alpha} l_\varepsilon$ of the boundary of $[-L, L]$. We can estimate the probability of this. On the set where the number of particles in the dual is at most $\varepsilon^{-2\alpha}$, this is

$$\leq C_2 \varepsilon^{-2\alpha} 8 \varepsilon^{-2\alpha} l_\varepsilon$$

for appropriate $C_2 > 0$. Combining this with (5.4) and (5.5) shows that

$$P(\zeta_T^\varepsilon(x) = 1) \leq u_L(T, x) + C_2 \varepsilon^{-2\alpha} 8 \varepsilon^{-2\alpha} l_\varepsilon + \varepsilon^{\alpha} + \kappa \varepsilon^{1 - 4\alpha}.$$

We will now show that

$$P(\xi_T^\varepsilon(x) = 1) \leq P(\zeta_T^\varepsilon(x) = 1) + \varepsilon^{\alpha} + 2\kappa \varepsilon^{-4\alpha} l_\varepsilon.$$

We use the same argument as in Lemma 4.3. Observing that on the set where the number of particles in the dual are $\leq \varepsilon^{-2\alpha}$ and no collisions occur, the system behaves as if it landed on a product measure provided that all the parents live in different subintervals. It follows from (5.4) and (4.10) that the complement of the set in the last sentence has probability at most $\varepsilon^{\alpha} + 2\kappa\varepsilon^{-4\alpha} l_{\varepsilon}$. Combining the last two estimates and (5.1), it follows that

$$P(\xi_T^{\varepsilon}(x) = 1) \le C_0 \varepsilon^{\gamma_0} + 2\varepsilon^{\alpha} + \kappa\varepsilon^{1-4\alpha} + 2\kappa\varepsilon^{-4\alpha} l_{\varepsilon}.$$

If we choose $\alpha$ and $l_{\varepsilon}$ so that $2\varepsilon^{-4\alpha} l_{\varepsilon}$ goes to 0 as $\varepsilon \to 0$, then we can choose $0 < C_1$, $\gamma_1 < \infty$ so that the lemma follows. □

**Lemma 5.2.** *If $\xi_0^{\varepsilon}$ has empirical one-point density at most $\rho_u/4$ and empirical two-point density at most $(\rho_u/3)^2$ in $[-L, L]$ and if $L_0 \le L \le L_1$ then with probability at least $1 - C_{13}\varepsilon^{\gamma_{13}}$, $\xi_T^{\varepsilon}$ has empirical one-point density at most $\rho_u/4$ and empirical two-point density at most $(\rho_u/3)^2$ in $[-L - vT, L + vT]$.*

**Proof.** This is the analogue of Lemma 4.3 in Durrett and Neuhauser (1994). Since $\mathrm{var}(Y) \le EY^2$,

$$\mathrm{var}(\tilde{\xi}_T^{1,\varepsilon}(x)) \le (\varepsilon/l_{\varepsilon})^2 \left\{ \sum_{y} E[\xi_T^{\varepsilon}(y)]^2 + \sum_{y \ne z} E[\xi_T^{\varepsilon}(y)][\xi_T^{\varepsilon}(z)] \right\},$$

where the summation ranges over $x \le y, z < x + l_{\varepsilon}$. It follows from Lemma 5.1 that the first sum is at most $C_1 \varepsilon^{1+\gamma_1} l_{\varepsilon}^{-1}$. To estimate the second sum we compare $\xi_T^{\varepsilon}(y)$ with the process defined in Section 2 whose duals are independent branching random walks on $\mathbb{R}$. This process was defined shortly before Lemma 2.1 and denoted by $\zeta_t$. Its dual was denoted by $Y = \{Y_s^k : k \ge 1\}_{s \ge 0}$. $\zeta_t$ computes its states from $\Phi_L(x)$. $\xi_t^{\varepsilon} \le \zeta_t$ if the following events occur: (i) there are not too many particles at time $T$, (ii) no collisions occur during $[0, T]$, and (iii) all the offspring produced in $X_{\varepsilon}$ (defined in Section 2) sample from sites that can be treated independently. Lemma 4.2 in Durrett and Neuhauser (1994) implies

$$P(\text{more than } \varepsilon^{-2\alpha} \text{ particles in the dual}) \le C_3 \varepsilon^{40\alpha}.$$

The probability of a collision in one dual or between the two duals before time $T$ when both have less than $\varepsilon^{-2\alpha}$ particles is smaller than $4\kappa\varepsilon^{-4\alpha+1}$ using (2.4). If we have less than $\varepsilon^{-2\alpha}$ particles in both duals then $X_{\varepsilon}^k$ and $Y^k$ are at most $2\varepsilon^{-2\alpha} l_{\varepsilon}$ apart. We can choose $\alpha$ and $\varepsilon$ so that they are automatically close to each other.

$X_{\varepsilon}$ samples from sites which can be treated independently if all the pairs land in different subintervals. This has probability at most $C_4 \varepsilon^{\gamma_4}$ as in the proof of Lemma 5.1. Therefore this is not much worse than sampling from a product measure with density $\rho/3$ since the two-point empirical density is at most $(\rho/3)^2$. The probability that any of the complements of the events in (i)–(iii) occur can therefore be bounded by

$$\le \varepsilon^{-2\alpha}[4\varepsilon^{1-4\alpha}\kappa + 2C_3\varepsilon^{40\alpha} + C_4\varepsilon^{\gamma_4}].$$

Combining all the estimates, we see that

$$E\{\xi^\varepsilon_T(y)\}\{\xi^\varepsilon_T(z)\} \le C_5\varepsilon^{\gamma_5}.$$

Hence, $\mathrm{var}(\tilde{\xi}^{1,\varepsilon}_T(x)) \le C_6\varepsilon^{\gamma_6}$. Since $E(\tilde{\xi}^{1,\varepsilon}_T(x)) \le C_1\varepsilon^{\gamma_1}$ by Lemma 5.1, Chebyshev's inequality implies that $P(\tilde{\xi}^{1,\varepsilon}_T(x) > \rho_u/4) \le C_7\varepsilon^{\gamma_7}$. We are concerned about the density in at most $2(L+vT)l_\varepsilon^{-1}$ intervals. Therefore, the probability that something might go wrong in any of the intervals is

$$\le 2(L+vT)l_\varepsilon^{-1}C_7\varepsilon^{\gamma_7} \le C_8\varepsilon^{\gamma_8}$$

for $L \le L_1$ for appropriately chosen $L_1$.

It remains to show that the empirical two-point density is at most $(\rho_u/3)^2$ in $[-L-vT, L+vT]$. As in the proof of Lemma 4.3, we bound

$$\mathrm{var}(\tilde{\xi}^{2,\varepsilon}_T(x)) = (\varepsilon/l_\varepsilon)^2\Bigg\{\sum_{y\in x+[0,l_\varepsilon)} \mathrm{var}(\xi^\varepsilon_T(y)\xi^\varepsilon_T(y+\varepsilon))$$

$$+ \sum_{\substack{y\ne z\\ y,z\in x+[0,l_\varepsilon)}} \mathrm{cov}(\xi^\varepsilon_T(y)\xi^\varepsilon_T(y+\varepsilon), \xi^\varepsilon_T(z)\xi^\varepsilon_T(z+\varepsilon))\Bigg\}.$$

The second sum can be split into two parts. The first part contains those terms where not all of the points in $\{y, y+\varepsilon, z, z+\varepsilon\}$ are different. The second part contains the remaining terms. The first part is of order $O(l_\varepsilon/\varepsilon)$ and can be combined with the first sum. The remaining part is at most $d_\varepsilon$. Therefore, this is

$$\le (\varepsilon/l_\varepsilon)^2(l_\varepsilon/\varepsilon)C_9 + d_\varepsilon.$$

Using a similar argument as for the one-point empirical density, $d_\varepsilon$ can be made smaller than $C_{10}\varepsilon^{\gamma_{10}}$ for some $0 < C_{10}, \gamma_{10} < \infty$. An application of Chebyshev's inequality gives

$$P(\tilde{\xi}^{2,\varepsilon}_T(x) > (\rho_u/3)^2 \quad \text{for some } x \in l_\varepsilon\mathbb{Z} \cap [-L-vT, L+vT])$$

$$\le C_{11}2(L+vT)l_\varepsilon^{-1}[(\varepsilon/l_\varepsilon)C_9 + C_{10}\varepsilon^{\gamma_{10}}].$$

This can be made $\le C_{12}\varepsilon^{\gamma_{12}}$ by choosing $L$ and $l_\varepsilon$ appropriately. □

**Lemma 5.3.** *Let $\delta = A_0/2$. Suppose $L_0 \le L \le L_1$ and $0 < \sigma < 1$. If ther are fewer than $\varepsilon^{\sigma-1}$ particles in $[-L, L]$ at time 0, then with high probability there are fewer than $\varepsilon^{\sigma+\delta-1}$ particles in $[-L+A_1T, L-A_1T]$ at time $T$.*

**Proof.** For $L \le L_1$, we choose $A_1$ so that with high probability

$$I^{x,T}_\varepsilon(t) \subset (-L, L) \quad \text{for all } t \le T,\ x \in [-L+A_1T, L-A_1T].$$

When none of the duals escapes from $[-L, L]$ by time $T$, then we do not have to worry about particles coming in from the outside. To find out how many particles are left after $T$ units of time, we have to estimate the number of births and deaths that will occur during the first $T$ units of time. Throughout the first $T$ units of time, the probability of a site being occupied is $\leq C_{14}\varepsilon^{\sigma}$. In the average, each vacant site in $[-L+A_1T, L-A_1T]$ has $\lambda T$ chances of becoming occupied if it finds two parents. As we saw earlier, distinct sites are almost independent for $\varepsilon$ small. Therefore, the expected number of births by time $T$ is at most

$$2L\varepsilon^{-1}\lambda T\varepsilon^{1.9\sigma},$$

since there are at most $2L\varepsilon^{-1}$ sites in $[-L+A_1T, L-A_1T]$. The $\varepsilon^{-0.1\sigma}$ takes care of the fact that adjacent sites are only almost independent. Markov's inequality then implies

$$P(\text{more than } \tfrac{1}{2}\,\varepsilon^{\sigma+\delta-1} \text{ births}) \leq C_{15}\,\varepsilon^{0.9\sigma-\delta}.$$

To estimate the number of deaths occurring, note that the probability a particle lives for more than $T$ units of time is $e^{-T} = \varepsilon^{A_0}$ by our choice of $T$. So the expected number of survivors is at most $C_{16}\varepsilon^{\sigma+A_0-1}$. Using Markov's inequality again

$$P(\text{more than } \tfrac{1}{2}\,\varepsilon^{\sigma+\delta-1} \text{ survivors}) \leq C_{17}\varepsilon^{A_0-\delta}.$$

If we choose $\delta = A_0/2$, the result follows from the two estimates. □

We will use Lemma 5.3 to show that if the empirical density is of order $\varepsilon^{\sigma}$ in a sufficiently large interval centered at the origin, then after a finite number of iterations the number of particles in that interval goes rapidly down to zero.

Let there be fewer than $\varepsilon^{\sigma/2-1}$ particles in $[-ML, ML]$. Pick $J_0$ such that $J_0\delta > 1$. Using Lemma 5.3 repeatedly then shows that with high probability there won't be any particles in $[-ML+J_0A_1T, ML-J_0A_1T]$. During the first $R$ units of time of our construction in the box, we use the traveling wave result to extend the interval of low density. From the preceding paragraph it is clear that we have to choose $M$ such that $ML < L_1$ and $ML - J_0A_1T \geq 5L$. The choice of $M$ then fixes $R$. We can then prove Lemma 5.4.

**Lemma 5.4.** *Suppose $\xi_0^{\varepsilon}$ has empirical one-point density $\leq \rho_u/4$ and empirical two-point density $\leq (\rho_u/3)^2$ in $[-L, L]$. If $\varepsilon$ is small, we can find $R$ such that $\xi_R^{\varepsilon}$ has empirical density $\leq \varepsilon^{\sigma}$ in $[-ML, ML]$ where $M$ satisfies the above relation.*

**Proof.** Using Lemma 5.1 and 5.3 repeatedly shows that we can get

$$P(\xi_R^{\varepsilon}(x) = 1) \leq C_{18}(\varepsilon^{\gamma_1} + (R/T+1)C_{13}\varepsilon^{\gamma_{13}})$$

for all $x \in [-ML, ML]$. Chebyshev's inequality then shows that with high probability

$$\xi_R^{\varepsilon} \text{ has less than } \varepsilon^{\sigma/2-1} \text{ particles in } [-ML, ML]. \qquad \square$$

To prove Theorem 4 we will now compare with $4N$-dependent oriented site percolation on $\mathscr{L} = \{(x, n) \in \mathbb{Z}^2: x + n \text{ is even}\}$. The good configurations are now those which are empty. The boxes are the same as before. We call a box $B_{x,n}$ *good* if the configuration in $2xL + [-5L, 5L]$ at time $(n+1)T$ is empty and if all the duals that are involved in determining whether or not the configuration is good, stay inside $2xL + [-2NL, 2NL]$. We start at time $-T$ from an initial configuration where all sites are occupied. We wait until time 0. At time 0 we look for an empty box where we can start our construction. Let $V_n = \{(x,n) \in \mathscr{L}$: there are no particles in $2xL + [-2L, 2L]$ at time $nT\}$. If $x \in V_0$, let $W_n^x = \{y: (x,0) \to (y,n)\}$. Let $l_n^x = \inf W_n^x$ and $r_n^x = \sup W_n^x$. We have to show

$$l_n^x \to -\infty, \quad r_n^x \to \infty \quad \text{a.s. on } \Omega_x \equiv \{W_n^x \neq \emptyset \text{ for all } n\} \tag{5.6}$$

and the following lemma.

**Lemma 5.5.** *With probability* 1 *we can find an* $x \in V_0$ *with* $W_n^x \neq \emptyset$ *for all* $n$.

The proof of (5.6) is a straightforward generalization of results in Durrett (1984) and can be found in Durrett and Neuhauser (1991). The paper just cited also contains a proof of Lemma 5.5.

To show extinction we have to find a completely vacant cone which roughly grows linearly in time. We will show that we can find this cone inside of the region bounded by $l_n^x$ and $r_n^x$. Since the kernel $k(\cdot)$ may have infinite support, not all boxes inside the region bounded by $l_n^x$ and $r_n^x$ are necessarily good. A site in this region can become occupied by checking parents outside of this region which then results in a bad box inside the region.

Let $\bigcup_n \tilde{W}_n^x$ be a subset of $\bigcup_n W_n^x$ consisting of adjacent sites such that $\tilde{W}_n^x \bigcap \tilde{W}_{n+1}^x \neq \emptyset$ for $n \geq 0$. Let $L_n^x = \inf \tilde{W}_n^x$ and $R_n^x = \sup \tilde{W}_n^x$. We furthermore require that $L_n^x - 2$, $R_n^x + 2 \notin W_n^x$ for $n \geq 0$. The set $\bigcup_n \tilde{W}_n^x$ is then the desired vacant cone.

To finish the proof it suffices to show that $R_n^x$ has a positive drift. Since $k(\cdot)$ has at least exponentially decaying tails, the probability that any of the sites to the left of the box corresponding to $R_n^x$ checks parents to the right of this box during $T$ units of time can be bounded by

$$\sum_{k \geq 1} [kLMT\lambda e^{-\gamma kL} + P(\text{more than } kLMT \text{ Poisson arrivals in the box } B(R_n^x - 2k, n))],$$

which can be made smaller than $q$ by choosing $L$ and $M$ large enough. This means that with probability at least $1 - q$, $R_n^x$ behaves like $r_n^x$. With probability at most $q$, it falls back a random distance. The drift of $R_n^x$ is therefore at least

$$(1-q)/3 - \sum_{k \geq 1} qC_{19} k e^{-\gamma_{19} k} \geq \delta > 0$$

if $q$ is sufficiently small. $C_{19}$ and $\gamma_{19}$ are arbitrary constants. The factor $\frac{1}{3}$ comes from a lower bound on the drift of $r_n^x$. This bound follows from results in Durrett (1984).

## 6. Metastability

In this section we will prove Theorem 5. We will show that there are parameter values where the process eventually dies out but does not show this for a very long time (of order $e^{\gamma/\varepsilon}$) when looking at a fixed finite interval. More precisely, we will show that when $\lambda > 4$ and when starting from all sites occupied, the empirical density in the interval $(0, 1)$ will, with high probability, approach $\rho_s$ and will stay close to $\rho_s$ for at least $T$ units of time where $T$ is such that $\varepsilon \log T \to \gamma' > 0$, as $\varepsilon \to 0$. We will first describe the strategy on how to prove this before we go into details. We will investigate the empirical density in small subintervals of length $a$ where $a$ is fixed and does not depend on $\varepsilon$, and show that with probability close to 1, it will stay above $\rho_s - \delta$ for $T$ units of time. We will start at time 0 with a large interval of length $C_1 e^{\gamma_1/\varepsilon}$, divide this interval into smaller subintervals of length $a$ and show that, with probability close to 1, after $\tau$ units of time – except for subintervals on both sides of the boundaries – the empirical density in all other subintervals will be close to $\rho_s$ when initially all the small subintervals have empirical density close to $\rho_s$. Every time we iterate this we will lose subintervals on both boundaries due to the fact that we do not have any control over the configuration outside of the interval. But since the interval we started with is very large, we can repeat this often enough so that after $T$ units of time we will not have lost any subintervals in $(0, 1)$ and will find enough particles in $(0, 1)$. We will therefore need an estimate on how many subintervals at the boundaries we lose in each iteration. This will be a large deviation estimate on how fast information can spread in the system. The other ingredient we need is an estimate on how much the empirical density changes in a subinterval during one iteration. We will start with the first estimate. For this, let $\xi_t^{(-\infty,0]}$ denote the sexual reproduction process starting from a configuration in which all sites $x \le 0$ are occupied at time 0 and all sites $x > 0$ are vacant at time 0. Let $r_t = \sup\{x\colon \xi_t^{(-\infty,0]}(x) = 1\}$ be the location of the rightmost particle at time $t$.

**Lemma 6.1.** *Let $M \le C_2 \varepsilon^{-3}$. For a fixed time $\tau > 0$, the probability that $r_t > M$ for some $t \in [0, \tau]$ is at most $C_3 e^{-\gamma_3/\varepsilon}$.*

**Proof.** We will do a very conservative estimate. We do not allow any deaths and start with a configuration where every site to the left of 0 is occupied. We fix a constant $V$. If the first birth to the right of 0 occurs within distance $V$ of 0, then we simply fill all the sites within distance $V$ of 0. We can then repeat this for the next birth to the right of this completely occupied interval. If the distance of the newborn particle to the occupied interval is bigger than $V$, we stop and say a bad event has occurred. We denote the distance between the newborn particle and the occupied interval by $X$. Since the kernel $k(\cdot)$ has exponential tails,

$$P(X > V) \le C_4 e^{-\gamma_4 V}.$$

The total rate at which a birth to the right of the occupied interval occurs, is at most $2\lambda V\varepsilon^{-1}$. The factor 2 takes into account parents whose distance to the boundary of the occupied interval is bigger than $V$. We set $\alpha \equiv 2\lambda V\varepsilon^{-1}$. By time $\tau$ we have with high probability at most $2\alpha\tau$ birth attempts. If we denote the position of the right edge of the occupied interval at time $t$ by $R_t$, then

$$P(R_\tau > 2\alpha\tau V) \leq C_5 e^{-\gamma_5 \alpha\tau} + 2\alpha\tau C_4 e^{-\gamma_4 V}.$$

The first term is a bound on the probability of having more than $2\alpha\tau$ birth attempts by time $\tau$. $C_5$ and $\gamma_5$ are arbitrary positive constants. The second term is a bound on the probability that any of the jumps are bigger than $V$. If we choose $V = \gamma_6/\varepsilon$ then the above estimate can be made smaller than $C_3 e^{-\gamma_3/\varepsilon}$ for arbitrary positive constants $C_3$ and $\gamma_3$. If $M \equiv 2\alpha\tau V = 4\lambda\tau\gamma_6^2\varepsilon^{-3} = C_2\varepsilon^{-3}$ for $C_2 = 4\lambda\tau\gamma_6^2$, then the lemma follows. □

From Lemma 6.1 we can conclude that with probability close to 1, we lose at most an interval of length $M = 2\alpha\tau V = 4\lambda\tau\gamma_6^2\varepsilon^{-3}$ on either side in each iteration. If we start with an interval of length $C_1 e^{\gamma_1/\varepsilon}$ then we can allow $\lfloor C_1 e^{\gamma_1/\varepsilon}/(2M+1) \rfloor$ iterations to reach time $T = \tau C_1 e^{\gamma_1/\varepsilon}/(2M+1) = Ce^{\gamma/\varepsilon}$ for some constants $C, \gamma \in (0, \infty)$. At time $T$, we will, with probability close to 1, still have an interval left where the empirical density is close to $\rho_s$.

The next step is to compute the change in the density of occupied sites during one iteration step. Let $u(n\tau) \equiv u(x, n\tau) = P(\xi_{n\tau}(x) = 1)$ and suppose that this quantity does not depend on $x$. Furthermore, assume that $u(n\tau) = \rho_s - \delta$ for some $\delta > 0$ sufficiently small. We wish to compute the change in the density during $\tau$ units of time in each subinterval of length $a$. A generator calculation yields

$$u((n+1)\tau) - u(n\tau) \geq -\tau u(n\tau) + \lambda\tau(1 - u(n\tau))(u(n\tau) - \tau)^2(1-\tau)(1-\kappa a) + O(\tau^2).$$

$O(s)$ denotes a function $g(s)$ so that $g(s)/s \to 1$ as $s \to 0$. $\kappa$ was defined in Section 2, i.e., $\kappa = \sup_{x\in\mathbb{R}} k(x)$. The first term on the right-hand side computes the change in density due to deaths, the second term changes due to births. In computing births, we will not use any sites that will become vacant during the iteration step. Also, we will only count those as newborn particle which will not die within $\tau$ units of their birth. Furthermore, we will not use parents which die during the iteration. The last factor $(1 - \kappa a)$ takes into account that we will not use any particles in the interval as parents. Simplifying the right-hand side of the above equation gives for $\delta$ and $\tau$ sufficiently small

$$= -\tau[-u(n\tau) + \lambda(1 - u(n\tau))u^2(n\tau)(1-\kappa a)] + O(\tau^2).$$

If $u(n\tau) = \rho_s - \delta$ with $\delta > 0$ and small, we can choose $a$ and $\tau$ small enough so that

$$u((n+1)\tau) - u(n\tau) \geq C_7\delta\tau \tag{6.1}$$

for an arbitrary, positive constant $C_7$.

We will use this estimate in a large deviation estimate on the empirical density of small subintervals. Fix $\delta$, $\tau$ and $a$ small so that (6.1) holds. Suppose at time 0, all the subintervals in $(-0.5C_1 e^{\gamma_1/\varepsilon}, 0.5C_1 e^{\gamma_1/\varepsilon})$ have empirical density $\rho_s - \delta$ and that all possible configurations in a given subinterval have the same probability conditioned on the number of particles in the subinterval. In particular, this implies that neighboring sites are independent. We already saw that $u(\tau) - u(0) \geq C_7\delta\tau$. Since each of the occupied sites in a given subinterval has the same probability of dying within $\tau$ units of time and each vacant site has the same probability of attempting to become occupied, then as long as vacant sites do not choose their parents within the same subinterval (which we excluded in our estimates), we simply have a Markov process in the subinterval where all states are equally likely when conditioned on the number of particles in the subinterval. To count the number of occupied sites in a given subinterval, we need some notation. For simplicity we will assume that $a/\varepsilon$ is an integer and investigate the interval $(0, a]$. Let $\{X_i\}_{i \geq 1}$ be independent and identically distributed random variables with

$$X_i = \begin{cases} 1 & \text{if } \xi(i\varepsilon) = 1, \\ 0 & \text{otherwise}, \end{cases}$$

for $1 \leq i \leq a/\varepsilon$. Let $S_n = \sum_{i=1}^n X_i$, the number of occupied sites among the first $n$ sites in $(0, a]$. If $P(X_i = 1) = q = 1 - P(X_i = 0)$ and if $\mu < q$, then a large deviation estimate for Bernoulli random variables gives

$$P(S_n < n\mu) \leq e^{-n\gamma_7}$$

for some arbitrary $\gamma_7 > 0$. In our case, $n = a/\varepsilon$, $q = \rho_s - \delta + C_7\delta\tau$, and $\mu = \rho_s - \delta$. We call a subinterval *bad* if its empirical density falls below $\rho_s - \delta$. Then

$$P(\text{subinterval is bad}) \leq e^{-\gamma_8/\varepsilon}.$$

There are $C_1 e^{\gamma_1/\varepsilon}/a$ subintervals in $(-0.5C_1 e^{\gamma_1/\varepsilon}, 0.5C_1 e^{\gamma_1/\varepsilon})$. Hence,

$$P(\text{any of the subintervals are bad}) \leq (C_1/a)e^{-(\gamma_8-\gamma_1)/\varepsilon} \leq C_9 e^{-\gamma_9/\varepsilon} \tag{6.2}$$

for arbitrary positive constants $C_9$ and $\gamma_9$. Note that the number of subintervals we have to check, decreases over time due to the losses at the boundaries. Therefore, (6.2) also holds for later iterations. Combining now the two main estimates (6.1) and (6.2), shows that the probability that anything goes bad during one iteration is at most $C_{10}e^{-\gamma_{10}/\varepsilon}$. We have at most $C_1 e^{\gamma_1/\varepsilon}/(2M+1)$ iterations. Therefore, with probability at least $1 - C_1 C_{10} e^{-(\gamma_{10}-\gamma_1)/\varepsilon}/(2M+1)$, the empirical density will be at least $\rho_s - \delta$ in $[0, 1]$ by time $C_1 e^{\gamma_1/\varepsilon}$. Choosing $\gamma_1$ so that $\gamma_{10} - \gamma_1 \equiv \gamma > 0$, finishes the proof of Theorem 5.

**Acknowledgements**

The author wishes to thank Rick Durrett for suggesting the problem and for fruitful discussions. Furthermore, thanks to the referees for helpful comments and to Larry Gray for pointing out a mistake.

## References

J.-B. Baillon and M. Théra, About the equation $k * u^2 = u$, preprint (1990).

M. Bramson, Survival of nearest particle system with low birth rate, Ann. Probab. 17 (1989) 433–443.

M. Bramson and R. Durrett, A simple proof of the stability criterion of Gray and Griffeath, Probab. Theory Related Fields 80 (1988) 293–298.

H.-N. Chen, On the stability of a population growth model with sexual reproduction on $\mathbb{Z}^2$, Ann. Probab. 20 (1992) 232–285.

R. Durrett, An infinite particle system with additive interactions, Adv. in Appl. Probab. 11 (1979) 355–383.

R. Durrett, Oriented percolation in two dimensions, Ann. Probab. 12 (1984) 999–1040.

R. Durrett, Lecture Notes on Particle Systems and Percolation (Wadsworth, Pacific Grove, CA, 1988).

R. Durrett, A new method for proving the existence of phase transition, in: K. S. Alexander, J. C. Watkins, eds. Spatial Stochastic Processes: a festschrift in honor of Ted Harris on his seventieth birthday (Birkhäuser, Boston, 1991).

R. Durrett, Predator-prey systems, in: K. D. Elworthy and N. Ikeda, eds., Asymptotic Problems in Probability Theory: Stochastic Models and Diffusions on Fractals, Pitman Research Notes in Mathematics, Vol. 283 (Longman, Essex, UK, 1993) pp. 37–58.

R. Durrett and L. Gray, Some peculiar properties of a particle system with sexual reproduction, in: Stochastic spatial processes, Lecture Notes in Mathematics, Vol. 1212 (Springer, New York, 1986).

R. Durrett and C. Neuhauser, Epidemics with regrowth in $d = 2$, Ann. Appl. Probab. 1 (1991) 189–206.

R. Durrett and C. Neuhauser, Particle systems and reaction-diffusion equations, Ann. Probab. 22 (1994) 289–333.

L. Gray, Duality for general attractive spin systems with applications in one dimension, Ann. Probab. 14 (1986) 371–396.

T.E. Harris, Nearest-neighbor Markov interaction processes on multidimensional lattices. Adv. in Math. 9 (1972) 66–89.

T.M. Liggett, Interacting Particle Systems (Springer, New York, 1985).

E. Kamke, Differentialgleichungen-Lösungsmethoden und Lösungen, Band 1 (Chelsea, New York, 3rd ed., 1948).

C. Noble, Equilibrium behavior of the sexual reproduction process, Ph.D. Thesis Cornell University (1989).

C. Noble, Equilibrium behavior of the sexual reproduction process with rapid diffusion, Ann. Probab. 20 (1992).

G. Swindle, A mean field limit of the contact process with large range, Probab. Theory Related Fields 85 (1990) 261–282.

H. Weinberger, Long-time behavior of a class of biological models, SIAM J. Math. Anal. 13 (1982) 353–396.